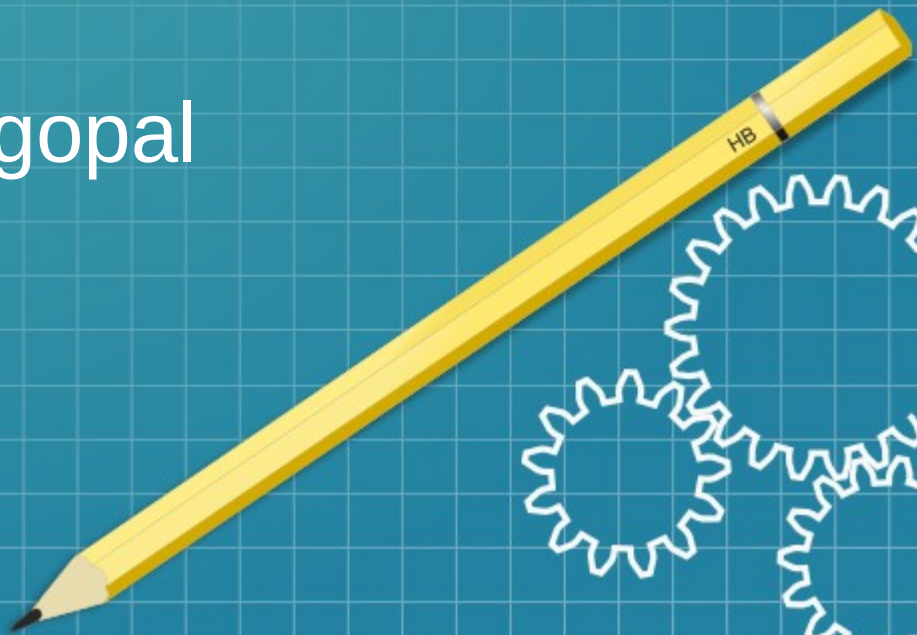
Capstone Project: The Battle of Neighborhoods

Ishwar Venugopal



Introduction

Consider an event management company named XYZ. A client approaches the company XYZ to organize a particular event in Toronto. As per the requirements of the client, the company is supposed to give a clear idea of the hotels and event spaces in the city along with the list of all Restaurants available in the vicinity of the hotels/event spaces. So as to facilitate this requirement, the company has to make a list of hotels and event spaces along with the Restaurants that are available nearby to these places.

Data

The data used in this project is provided by Foursquare location data. The data is sorted as per the venue categories that include hotels, event spaces and restaurants.

Methodology

- Using Foursquare account credentials, we retrieve the location data corresponding to the city of Toronto.
- We create a search query for Hotels and prepare the proper url for the same to get the data corresponding to Hotels

Methodology

- The json file obtained for Hotels is then converted into a pandas dataframe.
- Preprocessing is done on the pandas dataframe containing hotel data to get it in the required format with only the relevant data

	name	categories	address	lat	lng	postalCode	state
0	Sheraton Centre Toronto Hotel	Hotel	123 Queen Street West	43.651129	-79.383829	M5H 2M9	ON
3	DoubleTree by Hilton Hotel Toronto Downtown	Hotel	108 Chestnut Street	43.654608	-79.385942	M5G 1R3	ON
4	VFM Test Hotel	Hotel	123 Test Drive	43.658434	-79.387894	M2M 2M2	ON
7	Shangri-La Toronto	Hotel	188 University Ave.	43.649129	-79.386557	M5H 0A3	ON
9	Hilton Toronto	Hotel	145 Richmond St W	43.650143	-79.385488	M5H 2L2	ON

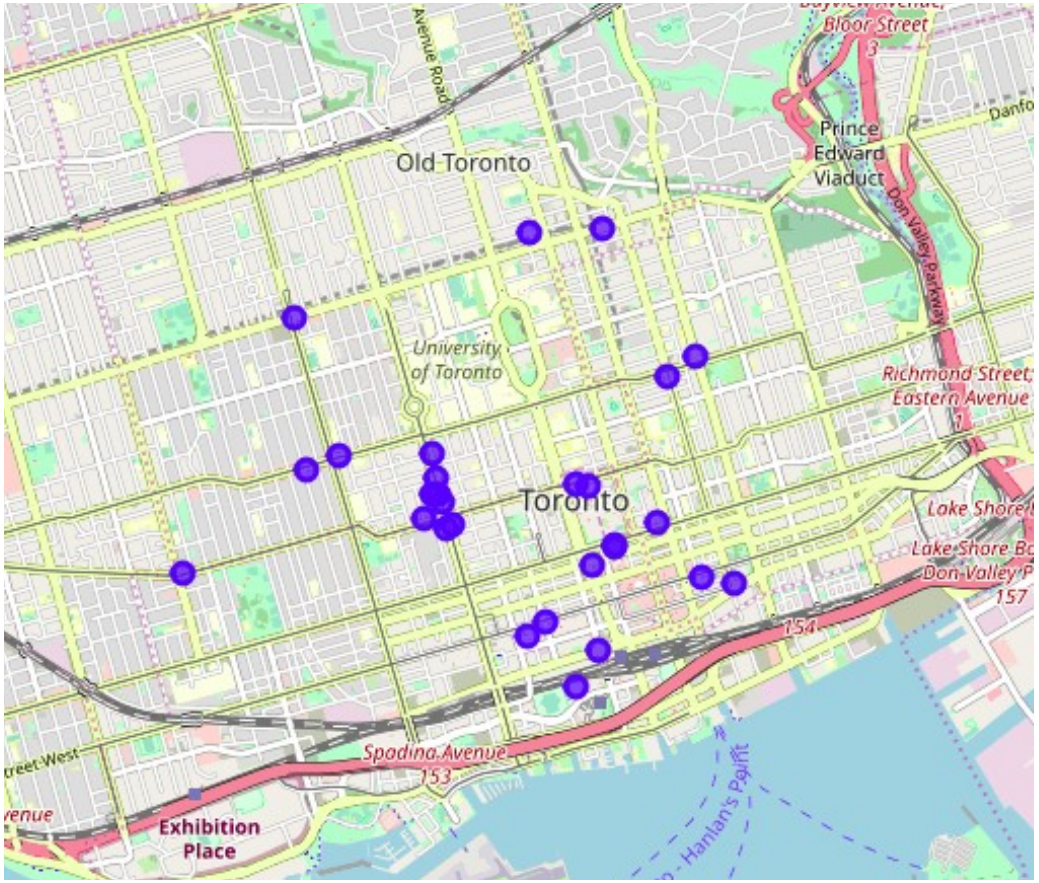
Methodology

- Now, we create a similar search query for Restaurants and prepare the proper url for the same to get the data corresponding to Restaurants
- The json file obtained for restaurants is then converted into a pandas dataframe.
- Preprocessing is done on the pandas dataframe containing Restaurant data similar to the method used earlier for Hotels, to get it in the required format with only the relevant data.

	name	categories	address	lat	lng	postalCode	state
0	Hemispheres Restaurant & Bistro	American Restaurant	110 Chestnut Street	43.654884	-79.385931	M5G 1R3	ON
1	Rol San Restaurant 龍笙棧	Dim Sum Restaurant	323 Spadina Ave.	43.654318	-79.398650	M5T 2E9	ON
2	Azure Restaurant & Bar	Restaurant	225 Frnt St W	43.644749	-79.385113	M5V 2X3	ON
3	Goldstone Noodle Restaurant 金石	Noodle House	266 Spadina Ave	43.652278	-79.398039	M5T 2E4	ON
4	Swatow Restaurant 汕頭小食家	Chinese Restaurant	309 Spadina Ave.	43.653866	-79.398334	M5T 2E6	ON

Results

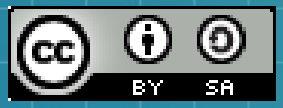
- A map showing these venues and how they cluster together is finally generated using the Folium library in Python.



Conclusions

- From the map, the client can now obtain areas corresponding to the regions which coincide with their choice of restaurants and hotels merging in a common area. This visualisation can also help the client in finalising a place for their event as per the different requirements of their visitors as well.

- The problem presented by the client was successfuly tackled and presented by the company XYZ by using Foursquare location data accompanied by Data Science tools and methodology through various packages/libraries in Python.



This project was created as part of the Applied Data Science Capstone Project on Coursera by Ishwar Venugopal.

Thank You!